OA Biochemistry
Open Access

Hollenstein M. Polymerize it all! A critical review. OA Biochemistry 2013 Jul 01;1(2):12.

Competing interests: none declared. Conflict of interests: none declared.
All authors contributed to conception and design, manuscript preparation, read and approved the final manuscript.
All authors abide by the Association for Medical Ethics (AME) ethical rules of disclosure.

Section: Enzymology

Polymerize it all! A critical review

Hollenstein, M.*

Department of Chemistry and Biochemistry, University of Bern, Freiestrasse 3, CH-3012 Bern, Switzerland.

Email: hollenstein@dcb.unibe.ch

Abstract

Introduction

Polymerases are fascinating and ubiquitous enzymes that catalyze the synthesis of nucleic acids with impressive rates and fidelity. Indeed, polymerases are capable of distinguishing minute differences in nucleoside triphosphates (e.g. the 2'-OH residue that is absent in dNTPs but present in rNTPs) and process with very low error rates (~10^{-6} for *Pfu* DNA polymerase, e.g.). However, this stringency is rather detrimental for practical applications since thermal stability (PCR), broad substrate tolerance (*in vitro* selections, functional tagging, etc.), or absence of exonuclease activity (sequencing) are desirable properties that are governed by the inherent nature of the polymerase. It is thus not surprising that polymerases are the targets of numerous engineering and modification strategies. This review highlights some of the recent progresses made in this field, with a particular emphasis on the substrate tolerance and processivity of mutant polymerases.

Conclusion

The development of new methods such as compartmentalized self-replication (CSR) selections and strategies for the screening of larger populations, has led to the generation of polymerases with tailor-made properties, capable of accepting a very broad range of substrates.

Keywords: Modified nucleic acids, polymerases, nucleoside triphosphates, PCR, protein engineering.

Introduction

For a long time, nucleic acids have been considered as function-deprived biopolymers that solely served as the blueprint of genetic information through their storage and replication abilities. However, the advent of solid-phase oligonucleotide synthesis and chemical genetics have strongly facilitated the appendage of exogenous chemical functionalities into nucleic acids with view on their use as therapeutic agents[1], building blocks for the expansion of the genetic alphabet[2], or for biosensing purposes[3]. In this context, suitably altered phosphoramidite units act as vectors for the selective introduction of functional groups into nucleic acids. However, the nature of the functionality is dictated by its resilience to the rather harsh synthetic conditions inherently imposed by the methodology itself.

On the other hand, the enzymatic polymerization of nucleoside triphosphates (dNTPs) bearing already built-in chemical groups has advanced as a versatile and alluring strategy for the modification of nucleic acids[4,5]. The advantage of modified dNTPs compared to the more static solid-phase synthesis is that large populations of oligonucleotides can be generated, which are then amenable to SELEX and related combinatorial methods of *in vitro* selection for the generation of functional nucleic acids[6-8]. Furthermore, polymerase catalyzed reactions with modified dNTPs allow to generate and explore sequences with vast chemical spaces. Indeed, the rather poor functional arsenal of DNA could be diversified by the incorporation of a myriad of modifications including bile acids[9], amino acids[10-12], norbornene residues[13], electrochemical tags[14], organocatalytic side chains[15], and even oligonucleotides[16]. The only restriction to this methodology is the substrate acceptance of the modified dNTPs by DNA polymerases. In this context, polymerases have evolved in nature to process with high fidelity and extremely low error rates (~10^{-6} for *Pfu* DNA polymerase)[17,18] and in the presence of modifications (either as a triphosphate or in the template strand) these enzymes will stall. Thus, the polymerization of modified dNTPs requires the use of engineered enzymes that overcome these natural traits. This review will examine recent progress made in the development and engineering of DNA and RNA polymerases with a special emphasis on substrate tolerance and processivity.

Discussion

Since most polymerases belong to superfamilies with similar sequences and structures, initial attempts at increasing their substrate acceptance and improving their replication properties focused on generating families of variants of the parent enzymes containing point mutations of residues[19]. This is usually achieved by sequence alignments followed by standard site-directed mutagenesis of the wild-type polymerase as highlighted in Figure 1 for the generation of the Sequenase DNA polymerase (*vide infra*).

Figure 1.

In particular, numerous polymerases that have been liberated of their 3'→5' or 5'→3' exonuclease proofreading activities were engineered. These *exo*$^-$ variants are more permissive to modified dNTPs while maintaining similar catalytic efficiencies (k_{cat}/K_m) to that of the parent wild-type polymerases. In this context, the deletion of 28 amino acids (Δ(Lys^{118}-Arg^{145}) of the native bacteriophage T7 DNA polymerase led to the isolation of a very potent polymerase, coined Sequenase (version 2.0), which kept the high activity of the parent enzyme while displaying a complete ablation of the exonuclease character[20]. In addition to high polymerization rates (~300 nucleotides per second), Sequenase presents a broader substrate tolerance and no exonuclease reactivity. Thus, before the advent of high-throughput and next generation sequencing (NGS) methods, Sequenase was a pivotal element in the Sanger sequencing strategy when used in conjunction with dideoxynucleoside triphosphates (ddNTPs)[21]. Moreover, Sequenase has proven to be more tolerant to rather arduous modifications such as 8-substituted dATP analogues (**1** in Figure 2), which are known to be poor substrates for most polymerases[22-25]. Similarly, the 9°N_m and Vent (*exo*$^-$) polymerases were obtained by single (Ala485→Leu) and double (Asp141→Ala and Glu143→Ala) mutations of the *Thermococcus* sp. 9°N-7 and *Tli* DNA polymerases, respectively[19,26,27]. Like Sequenase, 9°N_m and Vent (*exo*$^-$) polymerases are rather proficient at accepting and polymerizing modified dNTP substrates. Indeed, 9°N_m possesses a remarkable substrate tolerance since this enzyme is highly capable of incorporating a broad range of triphosphate analogues including the heavily, sugar-modified locked nucleic acid (LNA) **2** and α-L-

*threo*furanosyl nucleic acid (TNA) **3** derivatives (Figure 2) under primer extension reaction conditions[28-30]. The synthesis of TNA on a DNA template is less conspicuous when Vent (*exo*$^-$) polymerase is used since only one α-L-*threo*furanosyl building block can be incorporated into the nascent chain[31]. On the other hand, Vent (*exo*$^-$) polymerase displays a high tolerance for triphosphate analogues equipped with modifications anchored at the *N*7 and C5 positions of pyrimidines and of purines, respectively[11,12].

Figure 2.

Despite significant improvements in polymerase engineering and design, some unnatural nucleoside triphosphates still represent challenging hurdles that are difficult to overcome. In this context, improved screening methods could lead to the facile isolation of mutant polymerases, which in turn could alleviate the poor substrate uptake and extend the spectrum of modifications that can be polymerized efficiently and with high fidelity[32]. In this context, Marx *et al.* screened a library of mutants of the Therminator DNA polymerase for promoter-independent synthesis of RNA oligonucleotides[33]. The library, obtained by error-prone PCR and expression in 96-well plates[34], was then assessed for its ability to promote primer extension reactions with rNTPs. The most proficient mutant that was isolated could append up to 50 ribonucleotides on the DNA primer, leading to the generation of >70 nt long DNA-RNA hybrid sequences. In addition, this variant of the Therminator DNA polymerase could synthesize functional RNAs and incorporate 46 consecutive C5-modified dNTPs, features that are absent in the parent wild-type enzyme[33]. Furthermore, this mutant contained only a single Leu408→Gln mutation that seems to be crucial for substrate discrimination. This is an astounding finding, considering that natural polymerases need to differentiate dNTPs from rNTPs which are in a 10-100 fold excess in cells.

Another attractive and alternative method for the screening of large populations of polymerases to generate enzymes with tailor-made catalytic activities and substrate tolerances is the compartmentalized self-replication (CSR) directed evolution strategy[35]. Briefly, in a first step, a library of mutant polymerase genes is cloned and expressed in *E. coli* cells (Figure 3). The bacterial cells enclosing the polymerase and its encoding gene are supplemented with dNTPs and flanking primers prior to their compartmentalization in droplets of a water-in-oil

emulsion, which act as artificial cells and allow for both a physical separation of all the self-replicating systems and the association of phenotype and genotype of the same polymerase. The compartmentalized self-replicating units are then subjected to PCR amplification conditions. The first heat step causes the cell membranes to fissure and eventually disrupt, liberating the encoding gene and each individual polymerase, and initiating the replication process[35]. During the CSR selection, only active polymerases are able to self-replicate their encoded genes and will survive the selection pressure, much like in *in vitro* selections where only active nucleic acid enzymes proceed through the iterative Darwinian evolution process[36].

Figure 3.

Application of CSR led to the isolation of a *Taq* DNA polymerase variant that displayed an increased temperature resistance as compared to the wild-type enzyme. Indeed, clone T8 has a ~10 fold increase in half-life when subjected to a temperature of 97.5°C while maintaining similar catalytic efficiencies to that of wild-type *Taq*[35]. In addition, another clone, H15, displayed a dramatic resistance to the anticoagulant heparin as compared to the parent *Taq* DNA polymerase, albeit at the expense of a slight decrease in thermal stability. The CSR method also allowed for the selection of a *Taq* mutant that could process damaged DNA templates[37]. Indeed, the inclusion of mismatches (A·G or C·C) onto the flanking primers followed by three rounds of CSR selection afforded a polymerase variant that extended C·C mismatches with a >400-fold efficiency compared to the parent *Taq* DNA polymerase. This mutant also revealed proficient at extending templates that contained lesions such as abasic sites or thymidine dimers that usually cause polymerase-mediated reactions to stall[37]. More recently, in an impressive and elegant approach, a variant of CSR was pivotal for the crafting of polymerases that could accept non-natural analogues of DNA, both as triphosphates and as templates[38]. Indeed, in compartmentalized self-tagging (CST), only polymerases that are capable of recognizing modified dNTPs will extend biotinylated primers and form stable primer-plasmid duplexes that can then be captured with streptavidin-coated magnetic particles. The CST-based selection led to the isolation of mutants of the Tgo DNA polymerase, the replicative polymerase of *Thermococcus gorgonarius* that accept various sugar modified nucleosides such as TNA **3**, HNA **4**, CeNA **5**, and FANA **6** as substrates (Figure 2). Moreover, statistical correlation analysis followed by random mutagenesis led to

mutant polymerases with reverse transcriptase activity that recognized templates containing sugar-modified nucleotides. The synergic use of the engineered polymerases and reverse transcriptases in *in vitro* selection experiments culminated in the discovery of all-HNA aptamers against the HIV trans-activating response RNA motif and the hen egg lysozyme[38]. This further underscores the high potential and usefulness of CSR-based polymerase selection methods.

Finally, the development of alternative base pair systems in view of an expansion of the genetic code is an important research avenue[39]. In this context, the dZ:dP pair consisting of two non-standard nucleoside analogues (dZ **7** and dP **8**, Figure 2) with a different hydrogen bond pattern, has advanced as a potent surrogate of a natural base pair and is thus a potential candidate for a six-letter (ACGTZP) synthetic genetic system[40]. While polymerases such as *Taq* and Vent (*exo*$^-$) are rather tolerant to these nucleoside analogues under primer extension reaction and PCR conditions[41], the processivity and fidelity of their polymerization is still inferior to that of natural triphosphates[42]. In this context, a CSR selection was performed in order to generate a polymerase with higher incorporation rates of the non-standard triphosphate dZTP **7** opposite to dP **8** and form the orthogonal dZ:dP pair[42]. A nested PCR strategy with chimeric flanking primers containing embedded dP nucleotides was employed in the selection scheme. Screening of ~$50 \cdot 10^6$ *E coli* cells containing potential polymerases during the CSR selection yielded a *Taq* mutant that showed a higher propensity at synthesizing duplexes containing the dZ:dP pair. Surprisingly, the catalytic capacity of this *Taq* variant was completely ablated when the roles were inverted, i.e. when dPTP was to be incorporated opposite to dZ, clearly showing that, unlike natural DNA, these replication processes are not symmetrical.

Conclusion

Advances in molecular biology and protein engineering have boosted the development of DNA and RNA polymerases with tailored-made properties. Indeed, polymerases with increased heat-resistance, higher processivities, and broader substrate range have been crafted. For instance, it is now possible to generate rather long stretches (~75 nt) of constrained nucleic acids using selected polymerases, a task that is rather daunting when using standard

oligonucleotide synthesis. The generation of *de novo* polymerases with extended activities and substrate tolerance bodes well for the development of enzymes capable of polymerizing nucleic acid analogues that bear minimal similarities to their natural counterparts (e.g. peptide nucleic acids (PNA)) or even non-nucleosidic building blocks.

Acknowledgments

This work was supported by the Swiss National Science Foundation (grants no. PZ00P2_126430/1 and PZ00P2_144595). I would like to gratefully acknowledge Prof. Christian Leumann for providing me the lab space and equipment, as well as for his constant support. Dr. Pierre Mateo and Dr. Viviana Fluxá are acknowledged for their suggestions and for fruitful discussions.

Figure captions

Figure 1. Generation of polymerases with improved properties by site-directed mutagenesis. 1. Oligonucleotides containing the mutations at the appropriate position(s) are synthesized and annealed to the circular ssDNA containing the encoding gene (phage M13 mGP5-2, in the case of Sequenase, contains the sequence of the T7 gene 5 (shown in yellow) and upstream of the *lac* promoter (in green)); 2. The primer is then extended with a polymerase (native T7 DNA polymerase); 3. The resulting mutant circular ssDNAs are then transfected into *E. coli* cells where the polymerases will be expressed. After isolation and purification of the mutant polymerases, they can be evaluated for their activity[20,43].

Figure 2. Chemical structures of nucleoside analogues: 8-substituted dATP analogues **1** (R = e.g. (4-imidazoyl)-ethylamino-, 4-ethynylphenylalanine); locked nucleic acid (LNA) **2**; α-L-*threo*furanosyl nucleic acid (TNA) **3**; 1,5-anhydrohexitol nucleic acid (HNA) **4**; cyclohexenyl nucleic acid (CeNA) **5**; 2'-fluoro-arabinonucleic acid (FANA) **6**; dZTP **7**; dP **8**.

Figure 3. Schematic representation of the selection of polymerase by the compartmentalized self-replication (CSR) directed evolution method. 1. The initial library consists of *E coli* cells containing the encoding genes of the mutant polymerases (shown in red and green). The cells express the polymerase genes. 2. The bacterial cells containing the polymerases and their encoding genes are supplemented with flanking primer and dNTPs prior to being emulsified. 3. The first heat step causes the cellular membranes to break. 4. Only the polymerases (in green) that can operate under the given selection pressure conditions will self-replicate. Extraction and cloning yields an enriched library. 5. This library can then either be screened or used in a subsequent round of selection.

References

1. Wilson C, Keefe AD. Building oligonucleotide therapeutics using non-natural chemistries. Curr. Opin. Chem. Biol. 2006;10(6):607-614.
2. Krueger AT, Lu H, Lee AHF, Kool ET. Synthesis and Properties of Size-Expanded DNAs: Toward Designed, Functional Genetic Systems. Acc. Chem. Res. 2007;40:141-150.
3. Bell NM, Micklefield J. Chemical Modification of Oligonucleotides for Therapeutic, Bioanalytical and other Applications. ChemBioChem 2009;10:2691-2703.
4. Hocek M, Fojta M. Cross-coupling reactions of nucleoside triphosphates followed by polymerase incorporation. Construction and applications of base-functionalized nucleic acids. Org. Biomol. Chem. 2008;6(13):2233-2241.
5. Hollenstein M. Nucleoside Triphosphates - Building Blocks for the Modification of Nucleic Acids. Molecules 2012;17(11):13569-13591.
6. Franzen S. Expanding the catalytic repertoire of ribozymes and deoxyribozymes beyond RNA substrates. Curr. Opin. Mol. Ther. 2010;12:223-232.
7. Kuwahara M, Sugimoto N. Molecular Evolution of Functional Nucleic Acids with Chemical Modifications. Molecules 2010;15(8):5423-5444.
8. Mayer G. The Chemical Biology of Aptamers. Angew. Chem. Int. Ed. 2009;48:2672-2689.

9. Ikonen S, Macíčková-Cahová H, Pohl R, Šanda M, Hocek M. Synthesis of nucleoside and nucleotide conjugates of bile acids, and polymerase construction of bile acid-functionalized DNA. Org. Biomol. Chem. 2010;8(5):1194-1201.
10. Sakthivel K, Barbas CF. Expanding the Potential of DNA for Binding and Catalysis: Highly Functionalized dUTP Derivatives That Are Substrates for Thermostable DNA Polymerases. Angew. Chem. Int. Ed. 1998;37(20):2872-2875.
11. Jäger S, Rasched G, Kornreich-Leshem H, Engeser M, Thum O, Famulok M. A versatile toolbox for variable DNA functionalization at high density. J. Am. Chem. Soc. 2005;127(43):15071-15082.
12. Hollenstein M. Deoxynucleoside triphosphates bearing histamine, carboxylic acid, and hydroxyl residues – Synthesis and biochemical characterization. Org. Biomol. Chem. 2013;11:5162-5172.
13. Schoch J, Jäschke A. Synthesis and enzymatic incorporation of norbornenemodified nucleoside triphosphates for Diels–Alder bioconjugation. RSC Advances 2013;3(13):4181-4183.
14. Hocek M, Fojta M. Nucleobase modification as redox DNA labelling for electrochemical detection. Chem. Soc. Rev. 2011;40(12):5802-5814.
15. Hollenstein M. Synthesis of deoxynucleoside triphosphates that include proline, urea, or sulfamide groups and their polymerase incorporation into DNA. Chem. Eur. J. 2012;18(42):13320-13330.
16. Baccaro A, Steck A-L, Marx A. Barcoded Nucleotides. Angew. Chem. Int. Ed. 2012;51(1):254-257.
17. Cline J, Braman JC, Hogrefe HH. PCR fidelity of *Pfu* DNA polymerase and other thermostable DNA polymerases. Nucleic Acids Res. 1996;24(18):3546-3551.
18. Loakes D, Holliger P. Polymerase engineering: towards the encoded synthesis of unnatural biopolymers. Chem. Commun. 2009:4619-4631.
19. Gardner AF, Jack WE. Determinants of nucleotide sugar recognition in an archaeon DNA polymerase. Nucleic Acids Res. 1999;27(12):2545-2553.
20. Tabor S, Richardson CC. Selective Inactivation of the Exonuclease Activity of Bacteriophage T7 DNA Polymerase by in Vitro Mutagenesis. J. Biol. Chem. 1989;264(11):6447-6458.
21. Mardis ER. Next-Generation Sequencing Platforms. Annu. Rev. Anal. Chem. 2013;6:287-303.
22. Perrin DM, Garestier T, Hélène C. Expanding the catalytic repertoire of nucleic acid catalysts: Simultaneous incorporation of two modified deoxyribonucleoside triphosphates bearing ammonium and imidazolyl functionalities. Nucleosides Nucleotides 1999;18:377-391.
23. Čapek P, Cahová H, Pohl R, Hocek M, Gloeckner C, Marx A. An Efficient Method for the Construction of Functionalized DNA Bearing Amino Acid Groups through Cross-Coupling Reactions of Nucleoside Triphosphates Followed by Primer Extension or PCR. Chem. Eur. J. 2007;13:6196-6203.
24. Lam C, Hipolito C, Perrin DM. Synthesis and Enzymatic Incorporation of Modified Deoxyadenosine Triphosphates. Eur. J. Org. Chem. 2008:4915-4923.
25. Hollenstein M, Hipolito CJ, Lam CH, Perrin DM. A self-cleaving DNA enzyme modified with amines, guanidines and imidazoles operates independently of divalent metal cations (M^{2+}). Nucleic Acids Res. 2009;37(5):1638-1649.
26. Gardner AF, Jack WE. Acyclic and dideoxy terminator perferences denote divergent sugar recognition by archaeon and Taq DNA polymerases. Nucleic Acids Res. 2002;30(2):605-613.

27. Kong H, Kucera RB, Jack WE. Characterization of a DNA Polymerase from the Hyperthermophile Archaea *Thermococcus litoralis*. J. Biol. Chem. 1993;268(3):1965-1975.
28. Veedu RN, Vester B, Wengel J. Polymerase Chain Reaction and Transcription Using Locked Nucleic Acid Nucleotide Triphosphates. J. Am. Chem. Soc. 2008;130:8124-8125.
29. Horhota A, Zou K, Ichida JK, Yu B, McLaughlin LW, Szostak JW, Chaput JC. Kinetic Analysis of an Efficient DNA-Dependent TNA Polymerase. J. Am. Chem. Soc. 2005;127:7427-7434.
30. Yu H, Zhang S, Chaput JC. Darwinian Evolution of an Alternative Genetic System Provides Support for TNA as an RNA Progenitor. Nat. Chem. 2012;4(3):183-187.
31. Kempeneers V, Vastmans K, Rozenski J, Herdewijn P. Recognition of threosyl nucleotides by DNA and RNA polymerases. Nucleic Acids Res. 2003;31(21):6221-6226.
32. Kranaster R, Marx A. Engineered DNA Polymerases in Biotechnology. ChemBioChem 2010;11:2077-2084.
33. Staiger N, Marx A. A DNA Polymerase with Increased Reactivity for Ribonucleotides and C5-Modified Deoxyribonucleotides. ChemBioChem 2010;11:1963-1966.
34. Kranaster R, Marx A. Taking Fingerprints of DNA Polymerases: Multiplex Enzyme Profiling on DNA Arrays. Angew. Chem. Int. Ed. 2009;48:4625-4628.
35. Ghadessy FJ, Ong JL, Holliger P. Directed evolution of polymerase function by compartmentalized self-replication. Proc. Natl. Acad. Sci. USA 2001;98(8):4552-4557.
36. Joyce GF. Forty Years of In Vitro Evolution. Angew. Chem. Int. Ed. 2007;46:6420-6436.
37. Ghadessy FJ, Ramsay N, Boudsocq F, Loakes D, Brown A, Iwai S, Vaisman A, Woodgate R, Holliger P. Generic expansion of the substrate spectrum of a DNA polymerase by directed evolution. Nat. Biotechnol. 2004;22(6):755-759.
38. Pinheiro VB, Taylor AI, Cozens C, Abramov M, Renders M, Zhang S, Chaput JC, Wengel J, Peak-Chew S-Y, McLaughlin SH and others. Synthetic Genetic Polymers Capable of Heredity and Evolution. Science 2012;336:341-344.
39. Benner SA. Understanding Nucleic Acids Using Synthetic Chemistry. Acc. Chem. Res. 2004;37:784-797.
40. Yang Z, Chen F, Alvarado JB, Benner SA. Amplification, Mutation, and Sequencing of a Six-Letter Synthetic Genetic System. J. Am. Chem. Soc. 2011;133:15105-15112.
41. Yang Z, Sismour AM, Sheng P, Puskar NL, Benner SA. Enzymatic incorporation of a third nucleobase pair. Nucleic Acids Res. 2007;35:4238-4249.
42. Laos R, Shaw R, Leal NA, Gaucher E, Benner SA. Directed Evolution of Polymerases To Accept Nucleotides with Nonstandard Hydrogen Bond Patterns. Biochemistry 2013;52:5288-5294.
43. Dutta S, Li Y, Johnson D, Dzantiev L, Richardson CC, Romano LJ, Ellenberger T. Crystal structures of 2-acetylaminofluorene and 2-aminofluorene in complex with T7 DNA polymerase reveal mechanisms of mutagenesis. Proc. Natl. Acad. Sci. USA 2004;101(46):16186-16191.

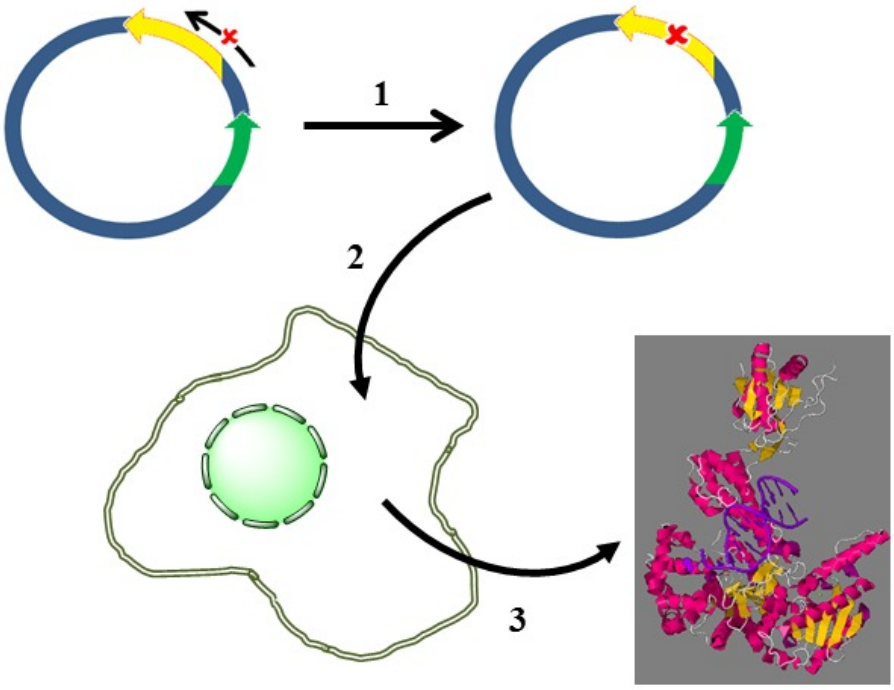

Figure 1: Figure1.tif

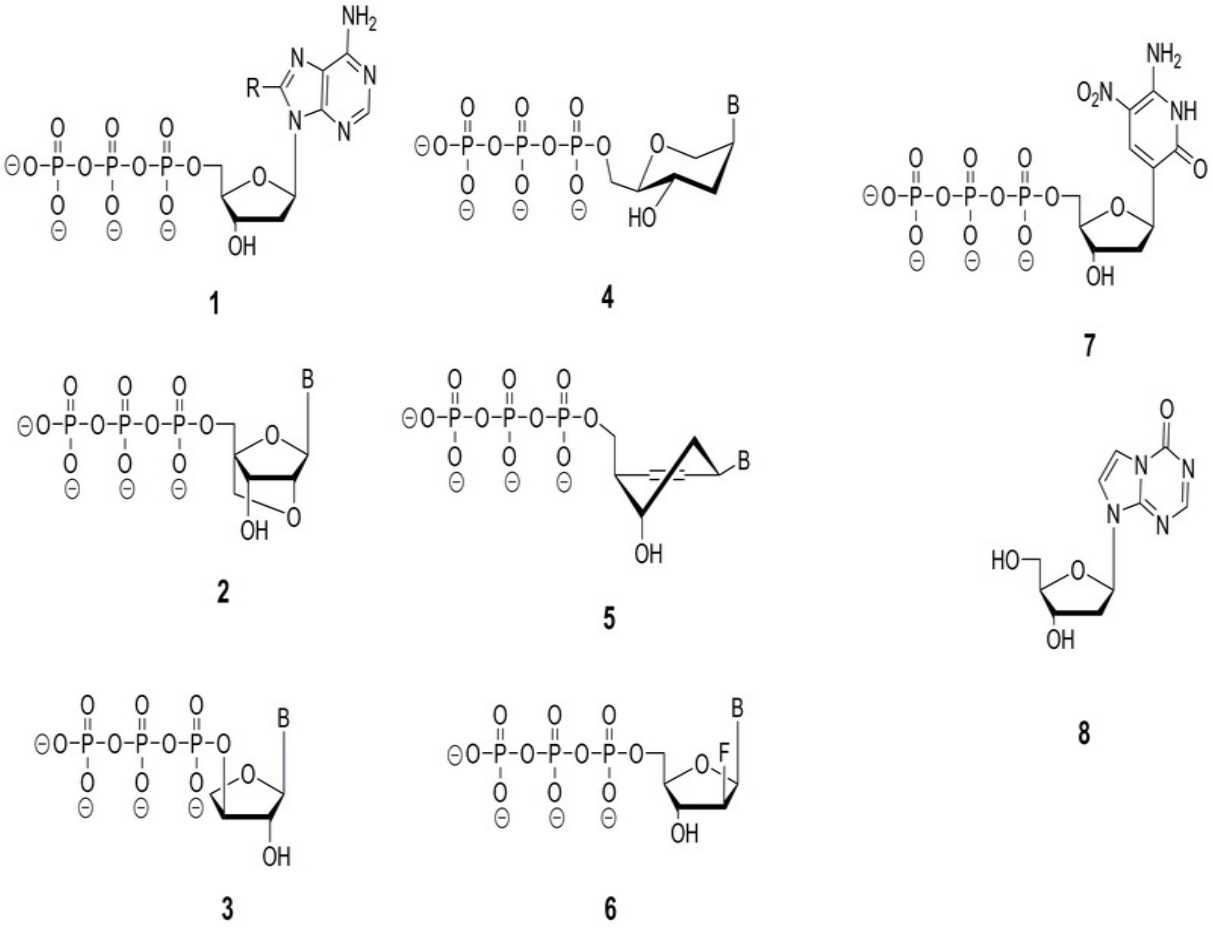

Figure 2: Figure2.tif

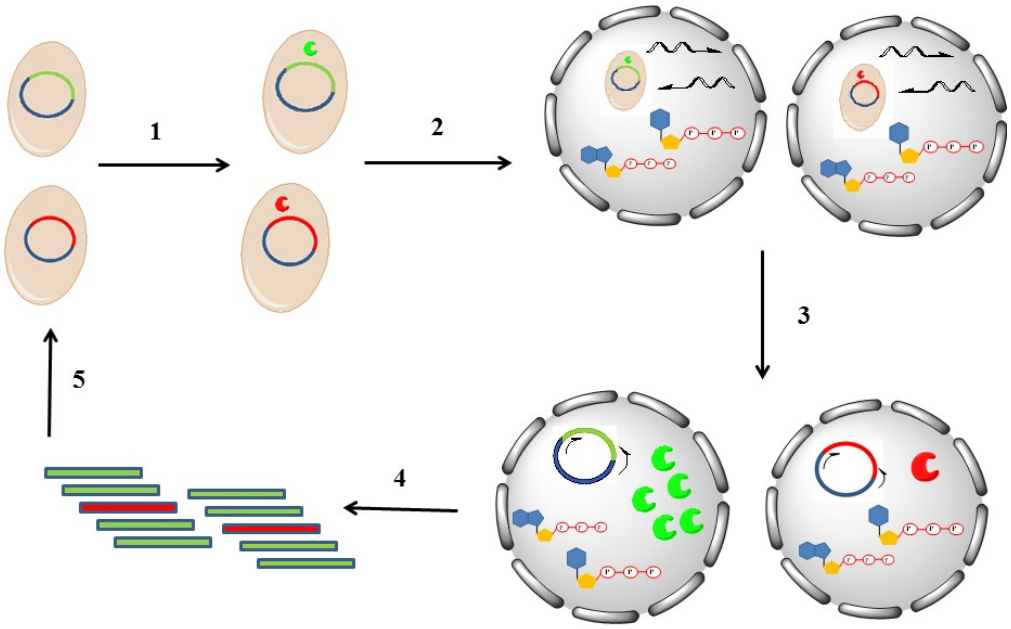

Figure 3: Figure3.tif